Treatment of a metabolic liver disease by in vivo genome base editing in adult mice

Villiger, Lukas ; Grisch-Chan, Hiu Man ; Lindsay, Helen ; Ringnalda, Femke ; Pogliano, Chiara B ; Allegri, Gabriella ; Fingerhut, Ralph ; Häberle, Johannes ; Matos, Joao ; Robinson, Mark D ; Thöny, Beat ; Schwank, Gerald

Abstract: CRISPR-Cas-based genome editing holds great promise for targeting genetic disorders, including inborn errors of hepatocyte metabolism. Precise correction of disease-causing mutations in adult tissues in vivo, however, is challenging. It requires repair of Cas9-induced double-stranded DNA (dsDNA) breaks by homology-directed mechanisms, which are highly inefficient in nondividing cells. Here we corrected the disease phenotype of adult phenylalanine hydroxylase (Pah) mice, a model for the human autosomal recessive liver disease phenylketonuria (PKU), using recently developed CRISPR-Cas-associated base editors. These systems enable conversion of C G to T A base pairs and vice versa, independent of dsDNA break formation and homology-directed repair (HDR). We engineered and validated an intein-split base editor, which allows splitting of the fusion protein into two parts, thereby circumventing the limited cargo capacity of adeno-associated virus (AAV) vectors. Intravenous injection of AAV-base editor systems resulted in Pah gene correction rates that restored physiological blood phenylalanine (L-Phe) levels below 120 µmol/l [5]. We observed mRNA correction rates up to 63%, restoration of phenylalanine hydroxylase (PAH) enzyme activity, and reversion of the light fur phenotype in Pah mice. Our findings suggest that targeting genetic diseases in vivo using AAV-mediated delivery of base-editing agents is feasible, demonstrating potential for therapeutic application.

DOI: https://doi.org/10.1038/s41591-018-0209-1

Originally published at:
Villiger, Lukas; Grisch-Chan, Hiu Man; Lindsay, Helen; Ringnalda, Femke; Pogliano, Chiara B; Allegri, Gabriella; Fingerhut, Ralph; Häberle, Johannes; Matos, Joao; Robinson, Mark D; Thöny, Beat; Schwank, Gerald (2018). Treatment of a metabolic liver disease by in vivo genome base editing in adult mice. Nature Medicine, 24(10):1519-1525.
DOI: https://doi.org/10.1038/s41591-018-0209-1

Treatment of a metabolic liver disease by *in vivo* genome base editing in adult mice.

Lukas Villiger[1], Hiu Man Grisch-Chan[7], Helen Lindsay[3,4], Femke Ringnalda[1], Chiara Balbo Pogliano[2], Gabriella Allegri[7], Ralph Fingerhut[7,8], Johannes Häberle[4,5,6], Joao Matos[2], Mark D. Robinson[3,4], Beat Thöny[5,6,7], and Gerald Schwank[1*]

[1]ETH Zurich, Department Biology, Institute for Molecular Health Sciences, Zurich, Switzerland

[2] ETH Zurich, Department Biology, Institute of Biochemistry, Zurich, Switzerland

[3]SIB Swiss Institute of Bioinformatics, University of Zurich, Zurich, Switzerland

[4]University of Zurich, Institute of Molecular Life Sciences

[5]Zurich Center for Integrative Human Physiology, Zurich, Switzerland

[6]Neuroscience Center Zurich, Zurich, Switzerland

[7]Division of Metabolism, University Children's Hospital Zurich and Children's Research Centre

[8]Swiss Newborn Screening Laboratory, University Children's Hospital Zurich, Switzerland.

*correspondence should be addressed to schwankg@ethz.ch (G.S.)

Abstract

CRISPR/Cas-based genome editing holds great promise for targeting genetic disorders including inborn errors of hepatocyte metabolism. Precise correction of disease-causing mutations in adult tissues *in vivo* however is challenging. It requires repair of Cas9-induced double stranded DNA (dsDNA) breaks by homology-directed mechanisms, which are highly inefficient in non-dividing cells. Here we corrected the disease phenotype of adult phenylalanine hydroxylase *(Pah)*^{enu2} mice, a model for the human autosomal recessive liver disease phenylketonuria (PKU)[1], using recently developed CRISPR/Cas-associated base editors[2–4]. These systems enable conversion of C·G to T·A base pairs and vice versa independent of dsDNA break formation and homology-directed repair (HDR). We engineered and validated an intein-split base editor, which allows splitting the fusion protein into two parts and thereby circumventing the limited cargo capacity of Adeno-associated virus (AAV) vectors. Intravenous injection of AAV-base editor systems resulted in *Pah*^{enu2} gene correction rates that restored physiological blood phenylalanine (L-Phe) levels below 120 µmol/l[5]. We observed mRNA correction rates up to 63%, restoration of phenylalanine hydroxylase (PAH) enzyme activity, and reversion of the light fur phenotype in *Pah*^{enu2} mice. Our findings suggest the feasibility of targeting genetic diseases *in vivo* using AAV-mediated delivery of base editing agents, demonstrating potential for therapeutic application.

Main text

Loss-of-function mutations in enzymes essential for hepatocyte metabolism constitute the majority of inborn liver diseases. Phenylketonuria is a widely studied autosomal recessive metabolic liver disease, where the deficiency of PAH enzyme activity leads to decreased metabolism of L-Phe, resulting in systemic hyperphenylalaninemia. Untreated infants affected by PKU suffer from severe retardation, microcephaly, and seizures[5–7]. Similar symptoms have been reported in untreated homozygous *Pah*^{enu2} mice, validating it as a model for human PKU[1,8,9]. *Pah*^{enu2} mice harbour a point mutation in the *Pah* gene on exon 7 (c.835T>C; p.Phe263Ser), which abolishes PAH function and causes abnormally elevated L-Phe levels above 1500 µmol/l[10]. Quantification of blood L-Phe levels offers a direct

readout for therapeutic efficacy, making the *Pah*^{enu2} mouse model a robust system to test gene therapy approaches[11]. Correction of the *Pah*^{enu2} disease model has previously been achieved by expressing functional enzymes from episomal cDNA templates[10–12]. Our approach, in contrast, aims to cure the disease by correcting the endogenous locus via genome editing.

Currently the most widely adopted systems for *in vivo* genome editing require Cas9 nuclease activity to introduce site-specific dsDNA breaks at targeted chromosomal loci. Several groups have previously applied this system *in vivo* to target disease-causing mutations in the mouse liver using exogenous DNA templates for HDR[13–15]. However, as HDR in non-dividing cells is highly inefficient compared to end-joining processes[16,17], the majority of dsDNA breaks generated in the adult liver resulted in random indel formations. Therapeutic application of CRISPR-associated nucleases to target genetic diseases in slowly proliferating tissues is therefore restricted to a small group of disorders, where either imprecise elimination of a gain-of-function mutation is sufficient[18,19], or the precise correction of a mutation confers a selective growth advantage to the edited cells[13,14]. Base editing is a strategy that allows genome editing independent of HDR and dsDNA break formation[3,4,20]. Base editors convert C·G to T·A base pairs or vice versa via mismatched U·G or I·A intermediates by either fusing a cytidine deaminase (rAPOBEC1) or an adenosine deaminase (TadA) to a catalytically dead Cas9 (dCas9)[3,4]. Importantly, these deaminases are single strand specific and editing is therefore limited to the protospacer region where the sgRNA/Cas9 complex binds to DNA[3,4,21]. Base editors have recently been employed in mice to knock out the *Pcsk9* gene in the liver in five week old mice, and to target Duchenne muscular dystrophy via intramuscular injection[22,23].
We reasoned that base editors allow precise correction of disease-causing mutations in non-dividing hepatocytes at rates sufficient to cure a disease phenotype. To test this hypothesis, we targeted the homozygous *Pah*^{enu2} c.835T>C mutation in the PKU disease model (Fig. 1a, b).

To specifically convert the c.835T>C mutation using cytidine deaminase base editors, we first searched for protospacer sequences that span the targeted base. We identified two protospacer adjacent motif (PAM) sites that allow binding of the corresponding Cas variants, *Staphylococcus aureus Sa*Cas9(KKH) (NNNRRT), and *Lacnospiraceae bacterium Lb*Cpf1(RR) (TYCV)[24–26] (Fig. 1a). An *Sa*Cas9(KKH) base editor (n*Sa*KKH-BE3) has

previously been established by Kim et al.[2]. We engineered a *Lb*Cpf1(RR) base editor, similar to *Lb*Cpf1 base editors recently described by Li et al[27]. In brief, we fused the cytidine deaminase rAPOBEC1 and an uracil glycosylase inhibitor (UGI) from the *Bacillus subtilis* bacteriophage PBS1[4] to a catalytically dead *Lb*Cpf1(RR) (referred to as d*Lb*RR-BE). We further constructed a *Lb*Cpf1 base editor with a mutated rAPOBEC1 cytidine deaminase (W90Y and R126E), which has been demonstrated to be less active and to narrow the activity window[2,27] (referred to as d*Lb*RR-minBE) (Fig. 1c, d). To evaluate editing efficiencies and activity windows of the base editors *in vitro,* we generated a reporter HEK293T cell line by stably integrating Exon 7 of the *Pah*^{enu2} allele. Transfection of plasmids expressing n*Sa*KKH-BE3, d*Lb*RR-BE, or dLbRR-minBE with the respective gRNAs showed editing efficiencies of 46%, 23.8%, and 4.2% for the target base at position 13, respectively (Fig. 1a, c, d, Suppl. Fig. 1). Komor et al. increased base editing efficiencies by using a (D10A) mutant of Cas9 (nCas9) to nick the target strand. As Cpf1 nickases that cleave the target strand have not been described, we attempted to increase editing efficiencies by co-transfecting a (D10A) nCas9 nickase. This nevertheless did not result in higher editing rates for d*Lb*RR-BE (Fig. 1c, d, Suppl. Fig. 1). Correction of the *Pah*^{enu2} locus critically depends on a suitable activity window, as conversion of two cytidines flanking c.835T>C lead to nonsynonymous mutations (Fig. 1a). We therefore developed a script to quantify the high-throughput sequencing (HTS) reads that restore the PAH amino acid sequence (Suppl. Fig. 1-4). Highest correction rates were observed with n*Sa*KKH-BE3, which was then selected for further studies (Fig. 1c, d, Suppl. Fig. 1).

To further assess the efficacy of the n*Sa*KKH-BE3 system, we next targeted the endogenous locus in primary liver cells. We established liver organoids lines from three *Pah*^{enu2} mice and transduced the n*Sa*KKH-BE3 system and the corresponding sgRNA using lentiviral vectors (Suppl. Fig. 5a). HTS of the targeted locus confirmed correction of the *Pah*^{enu2} allele in liver organoids, albeit at lower frequencies than in HEK293T reporter cells (Suppl. Fig. 5b).

Low immunogenicity and broad range of serotype specificity[28,29] suggest the use of AAV vectors for *in vivo* delivery of base editing agents, and prompted us to develop a strategy to circumvent the limited cargo capacity of AAV vectors (~4.8-4.9 kb, including inverted terminal repeats)[30]. We developed a dual AAV system and split n*Sa*KKH-BE3 in two parts[31], where each part is fused to the corresponding split-intein moiety from *Nostoc punctiforme (Npu)*[32] (referred to as p.N-int-BE3 and p.C-int-BE3.sgRNA) (Fig. 2a, b, Suppl. Table 1).

p.C-int-BE3.sgRNA co-expressed RFP and the sgRNA specific for *Pah*^{enu2}. We first confirmed association and functionality of both intein-split base editor moieties (Fig. 2c, d, Suppl. Fig. 5c, Suppl. Fig. 1). Despite lower abundance of the reconstituted base editor compared to the full-length version, editing rates were comparable (Fig. 2c, d, Suppl. Fig. 1). A similar observation has been made by Truong et al., who used the same intein system to split the *Sp*Cas9 nuclease[32]. Next we replaced the CMV promoter with the synthetic liver-specific promoter P3[11,33], and packaged the expression vectors into AAV2 serotype 8 particles (referred to as AAV8.N-int-BE3 and AAV8.C-int-BE3.sgRNA).

Adult mice between eight to ten weeks of age were injected with viral vectors (AAV8.N-int-BE3 and AAV8.C-int-BE3.sgRNA) in a one-to-one ratio via the tail vein (low dose = 5×10^{10} vector genomes (vg) per AAV per mouse, high dose = 5×10^{11} vg per AAV per mouse) (Fig. 2a). In control mice, the AAV8.C-int-BE3.sgRNA vector was replaced by a modified vector lacking the *Pah*^{enu2}-specific sgRNA (AAV8.C-int-BE3). Detection of RFP expression in the liver but not the pancreas confirmed liver-specific activity of the P3 promoter (Fig. 3a). The low-dose group reduced blood L-Phe levels to 732-1366 µmol/l eight weeks post injection (Fig. 3b). The high-dose group reduced blood L-Phe levels below therapeutic levels of 360 µmol/l[6] three to four weeks post injection (Fig. 3b). A further decrease of blood L-Phe levels to physiological levels below 120 µmol/l was observed six weeks post injection, representing a 20-fold reduction of initial L-Phe (Fig. 3b). Treated mice were followed up for 26 weeks and L-Phe levels remained below 120 µmol/l (Suppl. Fig. 6b). In addition, we observed comparable reduction in blood L-Phe when base editors were injected at higher viral titres (1×10^{12} vg per AAV) or at older age (12 weeks of age) (Suppl. Fig. 6c, d).

In a next step, we quantified *Pah* gene correction rates in mice treated for 4, 8, 14, and 26 weeks after injection of 5×10^{11} vg per AAV using HTS of PCR amplicons. We observe correction of 9.7% (6.1-12.7%) in mice after 4 weeks, 18.6% (8.8-28.5%) after 8 weeks, 22.1% (14.1-29.1%) after 14 weeks, and 25.1% (21.9-26.9%) after 26 weeks (Fig. 3c, d, e, Suppl. Fig. 2, 3). Indel mutations were detected in 4.0% (2.4-6.4%), 6.3% (1.1-11.6%), 9.6% (5.4-13.2%), 10.3% (10.0-10.5%), respectively (Fig. 3e, Suppl. Fig. 2, 3). Unedited reads were found in 82.5% (76.4-88.3%), 67.7% (51.0-85.7%), 60.7% (48.1-74.8%), and 56.8% (54.1-61.0%) in mice treated for 4, 8, 14, and 26 weeks (Fig. 3e, Suppl. Fig. 2, 3). Our data further show higher correction rates in male compared to female mice (Fig. 3d, Suppl. Fig. 2, 3), consistent with our observation that blood L-Phe levels decline more rapidly in males

(Suppl. Fig. 6e). These results could be explained by previously described gender-related differences of AAV transduction efficiencies in the liver[34].

Non-parenchymal cells account for 30-40% of total liver cells[35,36], where P3 promoter activity and AAV8 transduction efficiency is markedly reduced. We therefore reasoned that sequencing of genomic DNA from whole liver extracts underestimates c.835T>C correction rates in hepatocytes, and sequenced reverse-transcribed mRNA isolated from liver extracts. We found that *Pah*-mRNA correction rates were indeed higher, with 16.7% (13.3-20.8%) 4 weeks post injection, 34.4% (18.0-49.6%) 8 weeks post injection, 38.5% (26.1-47.6%) 14 weeks post injection, and 43.6% (39.1-47.1%) 26 weeks post injection (Fig. 3d, e). Strikingly, mice injected with a higher dose of 1×10^{12} vg per AAV showed mRNA correction rates up to 63% after 14 weeks (Suppl. Fig. 7a, 7b). As we observe a tendency of increased correction, both over time and with higher virus titres, we reason that low abundance of reconstituted base editor is likely a rate-limiting step in editing (Suppl. Fig. 5d, 7a, 7b). Interestingly, editing frequencies of different cytidines within the protospacer remain similar over time, indicating that long-term expression of base editors does not significantly affect the editing profile (Suppl. Fig. 7c).

We further quantified PAH enzyme activity in corrected mice 4 weeks and 8 weeks after administration of 5×10^{11} vg per AAV. Restoration of PAH enzyme activity (1.7-22.8% of wild-type enzyme activity) was confirmed in whole liver lysates and correlated to correction rates on mRNA and genomic DNA (Fig. 4a).
We next assessed reversion of PKU-associated phenotypes following gene correction. Reduced PAH activity in *Pah*^{enu2} mice limits melanin synthesis and manifests in hypopigmentation and a light fur phenotype[37]. In addition, homozygous *Pah*^{enu2} mice exhibit growth retardation, and are reduced in weight compared to heterozygous littermates[1]. Correction of the *Pah*^{enu2} allele led to a relative weight gain compared to untargeted homozygous *Pah*^{enu2} control mice (Fig. 4b), and changed the fur colour from agouti to black (Fig. 4c).

The clinical potential of CRISPR-associated base editors depends on their ability to limit editing to the targeted locus. Previous studies demonstrate that base editors do not induce untargeted C·G to T·A conversions randomly throughout the genome, but may cause undesired edits at regions where the sgRNA/base editor complex binds to DNA due to

sequence homology. We therefore identified ten potential off-target loci by computational prediction[38] (Suppl. Tab. 2,3) that were analysed in mice eight weeks after administration of 5×10^{11} vg per AAV. HTS revealed no C·G to T·A conversions or indel formations above background (Suppl. Fig. 8c). Furthermore, we find no indication of excessive DNA damage or cell proliferation after prolonged exposure to low levels of base editors (Suppl. Fig. 8a, b). Our data suggest that base editors in combination with highly specific guide RNAs have a low risk for generating off-target mutations, even when expressed over longer periods of time.

In conclusion, this work provides compelling evidence that AAV-mediated delivery of base editing agents allows to rescue the disease phenotype of the inborn metabolic liver disease PKU. We demonstrate proof of concept for *in vivo* gene correction using base editors in adult tissues with limited proliferative capacity. Messenger RNA correction rates up to 63% suggest applicability to a large number of genetic diseases.

Accession codes

High-throughput sequencing data have been deposited to EMBL-EBI Array Express (accession no. E-MTAB-7154. All other data are available from the corresponding author on reasonable request.

Code availability

Scripts for mapping sequencing data, counting editing events, and generating allele plots are available at https://github.com/HLindsay/Villiger_deaminase.

Data availability

All data and materials are available on reasonable request.

Acknowledgements

We thank Jean-Charles Paterna from the Viral Vector Facility of the Neuroscience Center Zürich for supplying vectors, the Functional Genomics Centre Zürich for sequencing in vivo samples, M. Rodriguez, A. Apladas, N. Rimann for their support with animal work, Andrea Garcete and Ines Kleiber from the Immunohistology laboratory for surgical research of the University Hospital Zurich for immunhistochemistry of liver samples.
This work was funded by the SNSF (31003A_160230). L.V. holds an MD/PhD scholarship from the Swiss National Science Foundation.

Author Contributions

L.V. designed the research, performed experiments, analysed data, and wrote the manuscript. F.R. conducted cryosections-, fluorescent imaging-, and western blot experiments and analysis thereof. J.M. and C.B.P performed western blot experiments. H.M.G. and G.A. assisted with *in vitro* PAH assays and data analysis thereof. H.L. wrote all R scripts and

assisted with HTS data analysis. R.F. analysed blood L-Phe levels. B.T., M.D.R., J.H., and H.M.G.-C. assisted with the design of experiments G.S. designed and supervised the research and wrote the manuscript. All authors approved the final version.

Competing financial interest

The authors declare no competing interests.

References

1. Shedlovsky, A., McDonald, J. D., Symula, D. & Dove, W. F. Mouse models of human phenylketonuria. *Genetics* **134,** 1205–1210 (1993).
2. Kim, Y. B. *et al.* Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat. Biotechnol.* **35,** 371–376 (2017).
3. Gaudelli, N. M. *et al.* Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* 1–27 (2017). doi:10.1038/nature24644
4. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533,** 420–424 (2016).
5. Blau, N., Van Spronsen, F. J. & Levy, H. L. Phenylketonuria. *Lancet* **376,** 1417–1427 (2010).
6. Mitchell, J. J., Trakadis, Y. J. & Scriver, C. R. Phenylalanine hydroxylase deficiency. *Genet. Med.* **13,** 697–707 (2011).
7. Scriver, C. R. & Clow, C. L. Epitome of human biochemical genetics. Part I. *new Engl.* **303,** 1394–1400 (1980).
8. Martynyuk, A. E. *et al.* Epilepsy in phenylketonuria: A complex dependence on serum phenylalanine levels. *Epilepsia* **48,** 1143–1150 (2007).
9. Martynyuk, A. E., van Spronsen, F. J. & Van der Zee, E. A. Animal models of brain dysfunction in phenylketonuria. *Mol. Genet. Metab.* **99,** S100–S105 (2009).
10. Ding, Z., Georgiev, P. & Thöny, B. Administration-route and gender-independent long-term therapeutic correction of phenylketonuria (PKU) in a mouse model by recombinant adeno-associated virus 8 pseudotyped vector-mediated gene transfer. *Gene Ther.* **13,** 587–

593 (2006).

11. Viecelli, H. M. *et al.* Treatment of phenylketonuria using minicircle-based naked-DNA gene transfer to murine liver. *Hepatology* **60,** 1035–1043 (2014).
12. Harding, C. O. *et al.* Complete correction of hyperphenylalaninemia following liver-directed, recombinant AAV2/8 vector-mediated gene therapy in murine phenylketonuria. *Gene Ther.* **13,** 457–462 (2006).
13. Yin, H. *et al.* letters Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. *Nat. Biotechnol.* **34,** 328–333 (2016).
14. Yin, H. *et al.* Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat. Biotechnol.* **32,** 551–3 (2014).
15. Yang, Y. *et al.* A dual AAV system enables the Cas9-mediated correction of a metabolic liver disease in newborn mice. *Nat. Biotechnol.* **34,** 334–338 (2016).
16. Mao, Z., Bozzella, M., Seluanov, A. & Gorbunova, V. Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA Repair (Amst).* **7,** 1765–1771 (2008).
17. Mao, Z., Bozzella, M., Seluanov, A. & Gorbunova, V. DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle* **7,** 2902–2906 (2008).
18. Ma, H. *et al.* Correction of a pathogenic gene mutation in human embryos. *Nature* **548,** 413–419 (2017).
19. Gao, X. *et al.* Treatment of autosomal dominant hearing loss by in vivo delivery of genome editing agents. *Nature* (2017). doi:10.1038/nature25164
20. Nishida, K. *et al.* Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **102,** 553–563 (2016).
21. Nishimasu, H. *et al.* Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **156,** 935–949 (2014).
22. Chadwick, A. C., Wang, X. & Musunuru, K. In Vivo Base Editing of PCSK9 (Proprotein Convertase Subtilisin/Kexin Type 9) as a Therapeutic Alternative to Genome Editing. *Arterioscler. Thromb. Vasc. Biol.* **9,** ATVBAHA.117.309881 (2017).
23. Ryu, S.-M. *et al.* Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy. *Nat. Biotechnol.* 1–7 (2017). doi:10.1038/nbt.4148
24. Kleinstiver, B. P. *et al.* Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523,** 481–5 (2015).
25. Zetsche, B. *et al.* Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas

System. *Cell* **163,** 759–771 (2015).
26. Gao, L. *et al.* Engineered Cpf1 variants with altered PAM specificities. *Nat. Biotechnol.* **35,** 789–792 (2017).
27. Li, X. *et al.* Base editing with a Cpf1-cytidine deaminase fusion. *Nat. Biotechnol.* **36,** 324–327 (2018).
28. Mingozzi, F. & High, K. A. Therapeutic in vivo gene transfer for genetic disease using AAV: Progress and challenges. *Nat. Rev. Genet.* **12,** 341–355 (2011).
29. Kay, M. A. State-of-the-art gene-based therapies: The road ahead. *Nat. Rev. Genet.* **12,** 316–328 (2011).
30. Ran, F. A. *et al.* In vivo genome editing using Staphylococcus aureus Cas9. *Nature* **520,** 186–190 (2015).
31. Nishimasu, H. *et al.* Crystal Structure of Staphylococcus aureus Cas9. *Cell* **162,** 1113–1126 (2015).
32. Truong, D. J. J. *et al.* Development of an intein-mediated split-Cas9 system for gene therapy. *Nucleic Acids Res.* **43,** 6450–6458 (2015).
33. Nair, N. *et al.* Computationally designed liver-specific transcriptional modules and hyperactive factor IX improve hepatic gene therapy. *Blood* **123,** 3195–3199 (2014).
34. Davidoff, A. M., Ng, C. Y. C., Zhou, J., Spence, Y. & Nathwani, A. C. Sex significantly influences transduction of murine liver by recombinant adeno-associated viral vectors through an androgen-dependent pathway. *Blood* **102,** 480–488 (2003).
35. Gao, B., Jeong, W. Il & Tian, Z. Liver: An organ with predominant innate immunity. *Hepatology* **47,** 729–736 (2008).
36. Racanelli, V. & Rehermann, B. The liver as an immunological organ. *Hepatology* **43,** (2006).
37. Thöny, B., Ding, Z., Rebuffat, A. & Viecelli, H. M. Phenotypic reversion of fair hair upon gene therapy of the phenylketonuria mice. *Hum. Gene Ther.* **25,** 573–4 (2014).
38. Haeussler, M. *et al.* Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* **17,** 1–12 (2016).

Figure legends

Figure 1 Concept to target the mutant *Pah^{enu2}* allele by genome base editing and *in vitro* validation in cell culture systems. a) Two gRNAs were designed to target the mutant *Pah^{enu2}* allele that harbours the disease-causing c.835T>C (p.Phe263Ser) mutation

on Exon 7, indicated in green. PAM sites are indicated in blue. One gRNA (sgRNA) allows binding of the *Sa*Cas9(KKH) variant to NNNRRT PAM sites, the other (crRNA) binding of the *Lb*Cpf1(RR) variant to TYCV PAM sites[26,46]. Different positions of cytidines amenable to deamination are numbered within the protospacer sequence. C·G to T·A conversion of the target C_{13} leads to the desired Ser > Phe change at amino acid position 263 that restores PAH enzyme activity. C·G to T·A conversion of C_6 or C_{15} leads to generation of a stop codon or an undesired nonsynonymous His to Tyr amino acid exchange, respectively. C·G to T·A conversion of C_{11}, C_{14} or C_{17} leads to synonymous mutations that do not affect the amino acid sequence. **b)** Conceptual outline to correct the mutated *Pah*^{enu2} locus in different models. **c)** Editing efficiencies in different C positions in the protospacer region determine activity windows of n*Sa*KKH-BE3, d*Lb*(RR)-BE, and d*Lb*(RR)-minBE. Experiments were performed in reporter HEK293T cells that have the mutated Exon 7 of the *Pah*^{enu2} gene stably integrated. Values represent mean ± s.d. of three independent biological replicates performed on separate days. A two-way ANOVA with Turkey's multiple comparisons test was performed for statistical analysis, *P (nSaCas9-BE3 + sgRNA vs. dLbCpf1-BE + crRNA) = 0.0461, *P (nSaCas9-BE3 + sgRNA vs. dLbCpf1-minBE + crRNA) = 0.0449, ****P < 0.0001. **d)** Correctly edited reads support restoration of the correct *Pah* gene sequence (C·G to T·A conversion of the target C_{13}, including conversions that lead to synonymous mutations) in reporter HEK293T cells. Incorrectly edited reads combine nonsynonymous mutations, non-C·G to T·A conversions and C·G to T·A conversions at positions other than C_{13}. Values represent mean of three independent biological replicates performed on separate days ± s.d. Statitical analysis was performed using a two-way ANOVA with Turkey's multiple comparisons test, ****P < 0.0001.

Figure 2 AAV intein-split base editors. a) Schematic maps of vector genomes for AAV8-delivery. N-int-BE3 expresses APOBEC1 fused to the N-terminal part of n*Sa*Cas9, and C-int-BE3 co-expresses the C-terminal part of n*Sa*Cas9 fused to an uracil glycosylase inhibitor, RFP and the *Pah*^{enu2}-specific sgRNA. **b)** Depiction of the two intein-split base editor proteins forming the full-length base editor after protein trans-splicing. **c)** Sequencing reads that support restoration of the correct PAH amino acid sequence after transfection of the intein-split system and the full length *Sa*KKH-B3 in reporter HEK293T cells. Transfection of N- or C-terminal parts alone and full length nSaKKH-BE3 without sgRNA serve as controls. Values and errors represent mean ± s.d. of three independent

biological replicates performed on different days. Statistical analysis was performed using a two-tailed, unpaired t-test. **d)** Western blot analysis of co-transfected intein-split N- and C-terminal parts of n*Sa*KKH-BE3. The FLAG epitope is only detected on the C-terminal part of the base editor. Transfection of N- or C-terminal parts alone and full length n*Sa*KKH-BE3 serve as controls. Similar results were obtained for two additional individually repeated experiments.

Figure 3 *In vivo* base editing corrects the disease-causing *Pah*^{enu2} mutation and leads to a reduction of blood L-Phe to physiological levels. Unless otherwise noted, treatment groups were n = 4 (two males, two females). 8-10 week old mice were administered 5×10^{11} vg of AAV8.N-int-BE3, and 5×10^{11} vg of AAV.C-int-BE3.sgRNA or 5×10^{11} vg of AAV.C-int-BE3 (untargeted control). **a)** Representative mouse liver cryosections in a mouse eight weeks after administration of 5×10^{11} vg per AAV. Red channel, RFP, blue channel, DAPI. Controls are sections of the pancreas from the same mouse and untreated liver sections from a *Pah*^{enu2} littermate. Scale bar, 100 µm. Similar results were obtained in one other independent experiment. **b)** Blood L-Phe levels from homozygous (targeted and untargeted) and heterozygous *Pah*^{enu2} mice were determined in weekly intervals. Heterozygous *Pah*^{enu2} littermates have physiological blood L-Phe levels and were used as controls. Values (n = 4 per treatment group) are presented as mean ± s.d., a two-way ANOVA with Dunnett's multiple comparisons test was performed to account for multiple comparisons to a single control (untargeted): *P = 0.0489, ****P < 0.0001. **c)** Editing efficiencies in different C positions within the protospacer region *in vivo* determine the activity window for mice after AAV injection. HTS was performed 4, 8, 14, and 26 weeks after injection of 5×10^{11} vg per AAV. Values (n = 4 for 4 and 8 weeks and n = 3 for 14 and 26 weeks) represent mean ± s.d. Statistical analysis was performed using a two-way ANOVA with Turkey's multiple comparisons test. **P (8 weeks treated vs. 26 weeks treated) = 0.0056, **P (14 weeks treated vs. 26 weeks treated) = 0.0095, ****P < 0.0001 **d)** Sequencing reads of genomic DNA and mRNA that support *Pah* gene restoration in mice 4, 8, 14, and 26 weeks after injection of 5×10^{11} vg per AAV. Each mouse is plotted separately. Horizontal bars represent mean values. **e)** Sequencing reads of genomic DNA and mRNA in untreated mice and mice 4, 8, 14, and 26 weeks after injection. Incorrectly edited reads combine nonsynonymous mutations, non-C·G to T·A conversions and C·G to T·A conversions at positions other than C_{13}. Values represent mean (n = 4 for 4 and 8 weeks and n = 3 for 14 and 26 weeks) ± s.d. **f)** Sequencing data of computationally

predicted off-target loci. Cumulative C·G to T·A conversions within the protospacer region were plotted for each off-target locus in (n = 4) mice 8 weeks after injection with 5×10^{11} vg per AAV, and compared to untreated mice (n = 4). Values are presented as mean ± s.d. Statistical analysis was performed using two-tailed, multiple t-test analysis.

Figure 4 In vivo base editing reliably rescues the disease phenotype in *Pah*^{enu2} mice. a) PAH enzyme activity in whole liver lysates of mice administered 5×10^{11} vg per AAV was determined at 4 weeks (n = 4) and 8 weeks (n = 4). Blood L-Phe levels at the point of euthanasia were 119-172 µmol/l and 79-106 µmol/l, respectively. PAH enzyme activity is normalized to wild-type *C57BL/6* mice (n = 4). Values are presented as mean ± s.d. Statistical analysis was performed using Pearson correlation to calculate R = 0.9717 (two-sided ****P > 0.0001) for genomic DNA and R = 0.9719 (two sided ****P > 0.0001) for mRNA. No adjustments were made for multiple comparisons. **b)** Homozygous *Pah*^{enu2} mice exhibit growth retardation that leads to lower weight. Depicted is the relative weight gain of homozygous *Pah*^{enu2} mice (n = 4) injected with 5×10^{11} vg per AAV to correct the mutation compared to age-matched homozygous *Pah*^{enu2} control mice (n = 4) injected with 5×10^{11} vg per AAV lacking the sgRNA. Values are presented as mean ± s.d. A two-tailed, unpaired t-test was used for statistical analysis: *P = 0.0346. **c)** The light fur phenotype of *Pah*^{enu2} mice 8 weeks after injection of 5×10^{11} vg per AAV. Controls are untreated *Pah*^{enu2} mice and wild-type *C57BL/6* mice. All mice shown are females and between 18-20 weeks of age. Similar results were obtained in all mice with corrected L-Phe levels.

Methods

Cloning. Sequences of constructs and primers used in this work are listed in the Supplementary Information. PCR was performed using Q5 High-Fidelity DNA Polymerase (New England Biolabs). pJL-SaKKH-BE3 and pBK-YE1-BE3 were a gift from David Liu (Addgene plasmid #85170 and #85174). WN10151 was a gift from Ervin Welker (Addgene plasmid #80441). pX601-AAV-CMV::NLS-SaCas9-NLS-3xHA-bGHpA;U6::BsaI-sgRNA was a gift from Feng Zhang (Addgene plasmid #61591), pLenti CMV GFP Puro (658-5) was a gift from Eric Campeau & Paul Kaufman (Addgene plasmid #17448). pCMV-VSV-G was a gift from Bob Weinberg (Addgene plasmid #8454) and psPAX2 was a gift from Didier Trono (Addgene plasmid #12260).

Cell culture transfection protocol and genomic DNA preparation. HEK293T (ATCC CRL-3216) and Hepa1-6 (ATCC CRL-1830) cell lines were maintained in Dulbecco's Modified Eagle's Medium plus GlutaMax (Thermo Fisher Scientific), supplemented with 10% (v/v) fetal bovine serum (FBS) and 1× penicillin-streptomycin (Thermo Fisher Scientific) at 37°C and 5% CO_2. Cells were maintained at confluency below 90%. HEK293T and Hepa1-6 cells were seeded on 48-well cell culture plates (Greiner) and transfected using 1.5 µl of Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's protocol. 12-16 h after seeding, at approximately 70% confluency, 1 µg total plasmid DNA was transfected for full length base editors (500 ng BE, 250 ng sgRNA or 250 ng crRNA, 250 ng empty pcDNA3.1), intein-split base editors (500 ng each of two parts), full length Cpf1-based base editors plus additional Cas9 nickase (500 ng BE, 250 ng crRNA, 150 ng SpCas9 D10A nickase, 100 ng sgRNA). Unless otherwise noted, cells were incubated for 5 days and split into a 24-well cell culture plate (Greiner) 24 h post transfection with a media change on day 4. Genomic DNA was isolated on day 5 using the DNeasy Blood and Tissue kit (Qiagen) according to the manufacturer's instructions.

Mouse liver organoid isolation and culture. Mouse liver tissue was mechanically dissociated and digested in Dulbecco's Modified Eagle's medium supplemented with Collagenase type IX 0.012%, Dispase 0.012%, and 1% (v/v) FBS for 30 min. Biliary ducts were identified in the microscope and seeded in 20µl Matrigel (Corning) drops. After gelation at 37°C, culture medium was added: Advanced Dulbecco's Modified Eagle's medium supplemented with 10% (v/v) RSPO1-conditioned medium prepared according to Farin et al.[39], 1 × B27 minus Vitamin A (Thermo Fisher Scientific), 1 × N-2 (Thermo Fisher Scientific), 10 mM Nicotinamine (Sigma-Aldrich), 1.25 mM N-acetylcysteine (Sigma-Aldrich), 10 nM gastrin (Tocris Bioscience), 10 nM Gastrin (Tocris Bioscience), 50 ng/ml EGF (Preprotech), 50 ng/ml HGF (Preprotech) and 50 ng/ml FGF (Preprotech). During the first 4-5 days after isolation, organoid culture medium was supplemented with 10% (v/v) Noggin- and 50% (v/v) Wnt3a-conditioned medium prepared according to Barker et al.[40]. Organoids were passaged by mechanical dissociation and transferred into fresh Matrigel drops.

AAV vector production. All pseudotyped AAV2/8 vectors were produced by the Viral Vector Facility of the Neuroscience Center Zurich. AAV vectors were ultracentrifuged and

diafiltered. Physical titers (vector genomes ml^{-1}) were determined using a Qubit® 3.0 Fluorometer. Identity of the packaged genomes of each AAV vector was confirmed by Sanger DNA-sequencing.

Lentiviral vector production. HEK293T cells were seeded in Opti-MEM (Thermo Fisher Scientific) in a T75 flask (Greiner) and transfected after 12-14 h at approximately 70% confluency using PEI. In brief, 59 µl PEI (0.1 mg/ml) in 370 µl Opti-MEM were incubated at room temperature for 5 minutes and added to 4.4 µg PAX2, 1.5 µg VSV-G and 5.9 µg lentiviral vector plasmid in 430 µl Opti-MEM. After incubation for 20 min at room temperature, cells were transfected. Media was changed 1 day after transfection. 2 days later, supernatant containing lentiviral particles was harvested and filtered using a Filtropur S 0.2 (Sarstedt) filter.

Lentiviral transduction of mouse liver organoids, HEK293T, and Hepa1-6 cells. Mouse liver organoids were dissociated into single cells using TrypLE (Thermo Fisher Scientific) at 37°C. After 5-6 min, three volumes of Dulbecco's Modified Eagle's Medium supplemented with GlutaMax (Thermo Fisher Scientific) and 10% (v/v) fetal bovine serum (FBS) were added, and samples were centrifuged. Cells were resuspended in supernatant containing lentiviral particles and Advanced Dulbecco's Modified Eagle's medium (Thermo Fisher Scientific) was added to a total of 500 µl (including lentiviral particles) and plated in 24 well suspension plates (Greiner). The plate was centrifuged for one hour at 32°C and 700 g, followed by a three-hour incubation at 37°C and 5% CO_2. Single cells were collected replated in 20 µl Matrigel (Corning) drops. Culture medium supplemented with 10µM Y-27632 (Tocris Bioscience) was added. After 2 days, stably transduced organoids were enriched with 3 µg/ml Puromycin (Invivogen) for 7 days. Supernatant containing lentiviral particles was added to HEK293T and Hepa1-6 cells in a 24 well cell culture plate (Greiner). Two days after transduction, cells were enriched using 2.5µg/ml Puromycin for seven days.

Animal studies. All animal experiments were performed in accordance with protocols approved by the Kantonales Veterinäramt Zürich in compliance with all relevant ethical regulations. *Pah^{enu2}* mice were housed in a pathogen-free animal facility at the Institute of Molecular Health Sciences at ETH Zurich and kept in a temperature- and humidity-controlled room on a 12h light-dark cycle. Mice were fasted for 3-4 hours before blood

was collected from the tail vein for L-Phe determination but otherwise were fed a standard laboratory chow (Kliba Nafag #3437 with 18.5% crude protein). Mice were genotyped at weaning. Heterozygous *Pah*^{enu2} littermates with comparable blood L-Phe levels to wild type *C57BL/6* mice (Suppl. Fig. 1d), were used as controls for physiological blood L-Phe levels (< 120 µmol/l). Wild-type control animals for PAH enzyme activity assays were age- and sex-matched *C57BL/6* mice purchased from JANVIER LABS. Unless otherwise noted, animals were allocated to groups of four and stratified by gender (2 males, 2 females). The low-dose group was injected with 5×10^{10} viral vector genomes (for each AAV) per mouse. The high-dose group was administered 5×10^{11} viral vector genomes (for each AAV) per mouse. Animals were euthanized at 4 weeks and at 8 weeks after injection for further analyses. The untargeted control group received the same vector dose as the high-dose group.

Amplification and high-throughput DNA sequencing of genomic DNA and mRNA samples.

Genomic DNA and mRNA from mouse liver tissue was isolated from whole liver lysate using the DNeasy Blood and Tissue kit or the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. cDNA was generated using GoScript reverse transcriptase (Promega) according to the manufacturer's protocol using Oligo-dT primers. Subsequent PCR reactions to generate amplicons for HTS were performed using NEBNext High-Fidelity 2x PCR Master Mix. In brief, 200 ng genomic DNA was amplified in 26 cycles for the first PCR in a 20 µl reaction. Similarly, cDNA from 50 ng reverse transcribed RNA was amplified in 22 cycles. The PCR product was purified using Agencourt AMPure XP beads (Beckman Coulter), and amplified with primers containing sequencing adaptors. The products were gel purified and quantified using the Qubit 3.0 fluorometer with the dsDNA HS assay kit (Thermo Fisher Scientific). Samples were sequenced on an Illumina Miseq.

HTS data analysis. Sequencing reads were demultiplexed using Miseq Reporter (Illumina), and analysed using a custom script. In short, reads were merged with PEAR v0.9.8[41] and mapped to the Ensembl mouse genome v38.90 using BWA MEM[42]. Base editing frequency was quantified in R using CrispRVariants v1.7.5[43] and Biostrings v2.46.0[44]. Scripts are available at https://github.com/HLindsay/Villiger_deaminase.

Mouse liver and pancreas cryosections. Mice were euthanized with CO_2 and the pancreas isolated immediately before perfusion through the portal vein using PBS followed by freshly prepared PLP buffer containing 75 mM L-Lysine (Sigma-Aldrich), 30.4 mM Na_2HPO_4, 7.1 mM NaH_2PO_4 (Sigma-Aldrich), $NaIO_4$ (Sigma Aldrich) and 1% PFA. The mouse liver and pancreas were fixed in PLP buffer overnight at 4°C, followed by 3 washing steps with buffer containing 81 mM Na_2HPO_4 and 19 mM NaH_2PO_4 at pH 7.4. Liver pieces were transferred to a 30% sucrose solution for 6 hours at 4°C and embedded in OCT compound in cryomolds (Tissue-Tek). Frozen liver tissue was sectioned at 10 µm at -18°C, and mounted directly on SuperFrost Plus slides (Thermo Fisher Scientific). Cryosetions were counterstained with DAPI (Thermo Fisher Scientific) and mounted in Vectashied mounting medium (Vector Labs). At least 3 frozen sections per mouse and tissue were analysed.

Microscopy. Mouse liver and pancreatic tissue was imaged using an inverted laser scanning microscope (Leica TCS SP8) in a XYZT-mode. To visualize DAPI and RFP, samples were excited at 405 nm, and 568 nm, respectively. Imaging conditions and intensity scales were matched for all images. Images were analysed using Leica LAS AF (Lite) software v3.3 and deconvoluted using Fiji ImageJ software (v1.51n).

Western blot analysis for FLAG epitope detection. HEK293T cells were transfected in a 48-well plate (Greiner) at 70% confluency as previously described and kept for 5 days. Cells were washed with PBS (Gibco) and harvested using RIPA lysis buffer (150 mM Tris pH 8.0, 150 mM NaCl, 0.1% SDS, 0.5% Na-Deoxycholate, 1% NP-40) supplemented with a protease inhibitor cocktail (Roche) without EDTA. Total protein was quantified using a Pierce Protein BCA assay kit (Thermo Fisher Scientific). Frozen mouse liver tissue was lysed in RIPA buffer (150 mM Tris pH 8.0, 150 mM NaCl, 0.1% SDS, 0.5% Na-Deoxycholate, 1% NP-40) supplemented with a protease inhibitor cocktail (Roche) without EDTA and treated analogously. 20 µg of total protein lysates were loaded and run on SDS-PAGE gels, transferred to nitrocellulose membranes (GE Life Sciences), and incubated with primary antibodies for FLAG (F1804 M2, Sigma-Aldrich, 1:1000) and actin (#4970, Cell Signaling Technology) over night at 4°C according the manufacturer's protocol. Membranes were washed three times with TBS with 0.1% Tween-20 (TBS-T) for 10 min and labelled with secondary anti-IgG-HRP antibodies raised against each corresponding primary antibody. After three washes with TBS-T, membranes were incubated with ECL

chemiluminescent reagent (Biorad) according the manufacturer's instructions and exposed using a FUSION SOLO S (Vilber). Uncropped versions of images are provided in the Supplementary information.

Western blot analysis for γ-H2AX-S139-P and KAP1- S824-P detection. Cell lysates were subjected to NuPAGE gel electrophoresis (NuPAGE 3-8% Tris-Acetate; NuPAGE 10% Bis-Tris, Invitrogen) and blotted on PVDF membranes (GE Healthcare). Proteins were detected using Rabbit anti-γ-H2AX-S139-P (ab2893, Abcam) 1:1500, Rabbit Anti-KAP1-S824-P (ab70369, Abcam) 1:2000, Mouse Anti-GAPDH (#GA1R, ThermoFisher) 1:10000.

PAH enzyme activity assay. Whole liver extracts were analysed using isotope-dilution liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) according to a previously published method[45].

Statistical analyses. A priori power calculations to determine sample sizes for animal experiments were performed using the R 'pwr' package. Statistical analyses were performed using GraphPad Prism 6.01 for Windows. Sample sizes and statistical tests used are described in figure legends. In brief, the Dunnett's test was used to compare multiple variables to a single control for blood L-Phe levels. All tail vein injections were successful and no animals were excluded. Two-way ANOVA analysis was used to compare different conditions (different base editors or virus titres) and different C positions. A two-tailed, unpaired t-test was used for analysis of relative weight gain between two groups and correctly edited reads *in vitro* or in treated animals. A non-parametric, two-tailed correlation test was used to analyse genomic DNA correction and PAH enzyme activity, as well as genomic DNA and mRNA correction. Analyses of off-targets and gender-related difference in gene correction were performed using multiple t-tests. Group averages are presented as mean ± s.d.

Methods-only references

39. Farin, H. F., Van Es, J. H. & Clevers, H. Redundant sources of Wnt regulate intestinal stem cells and promote formation of paneth cells. *Gastroenterology* **143,** 1518–1529.e7

(2012).

40. Barker, N. *et al.* Lgr5+veStem Cells Drive Self-Renewal in the Stomach and Build Long-Lived Gastric Units In Vitro. *Cell Stem Cell* **6,** 25–36 (2010).
41. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30,** 614–620 (2014).
42. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. **00,** 1–3 (2013).
43. Lindsay, H. *et al.* CrispRVariants charts the mutation spectrum of genome engineering experiments. *Nat. Biotechnol.* **34,** 701–702 (2016).
44. Pagès H, Aboyoun P, G. R. and D. S. Biostrings: Efficient manipulation of biological strings.
45. Heintz, C., Troxler, H., Martinez, A., Thöny, B. & Blau, N. Quantification of phenylalanine hydroxylase activity by isotope-dilution liquid chromatography-electrospray ionization tandem mass spectrometry. *Mol. Genet. Metab.* **105,** 559–565 (2012).
46. Kleinstiver, B. P. *et al.* Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. *Nat. Biotechnol.* **33,** 1293–1298 (2015).

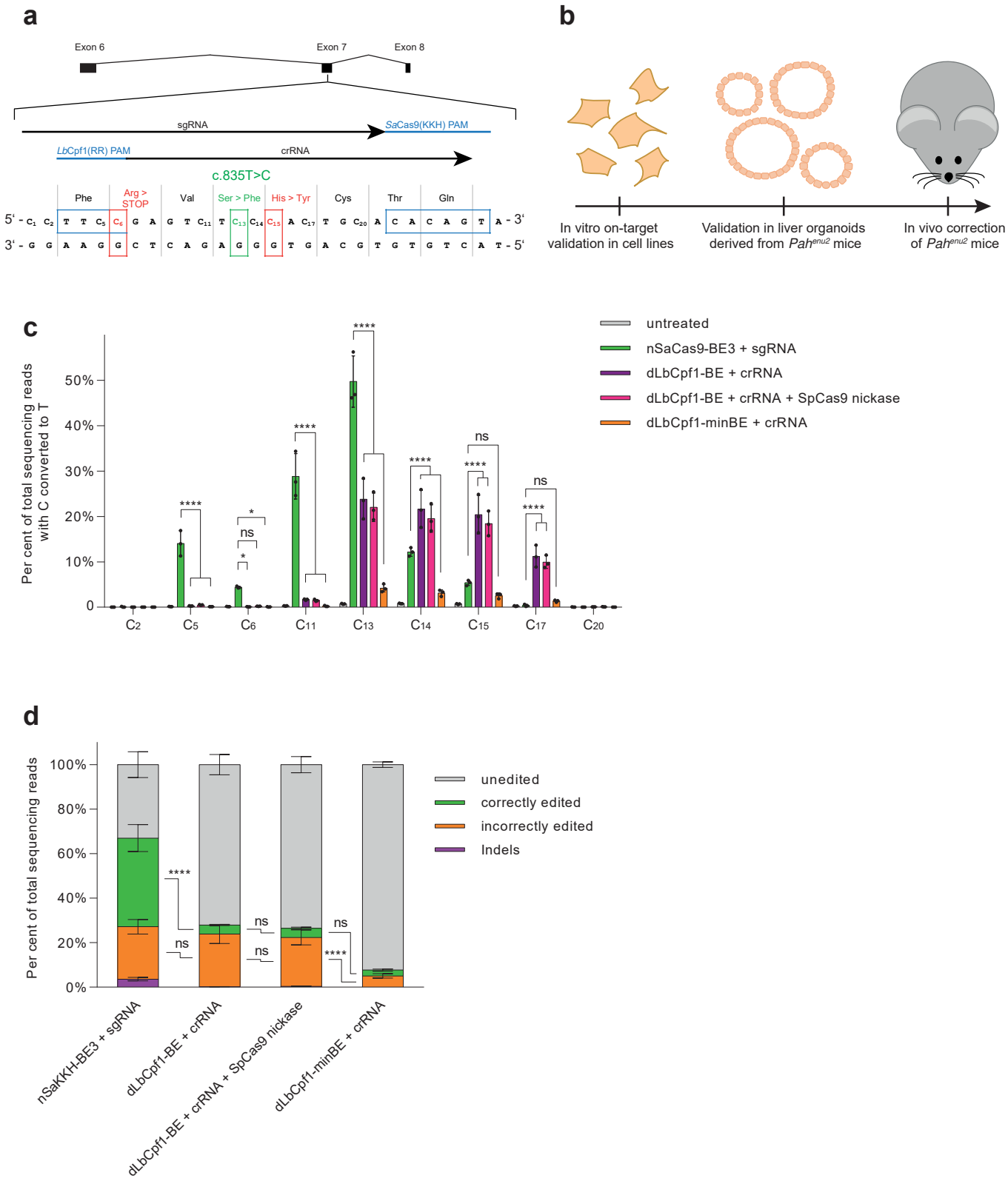

a
Exon 6
Exon 7
Exon 8
sgRNA
SaCas9(KKH) PAM
LbCpf1(RR) PAM
crRNA
c.835T>C
Phe
Arg > STOP
Val
Ser > Phe
His > Tyr
Cys
Thr
Gln
5' - C1 C2 T T C5 C6 G A G T C11 T C13 C14 C15 A C17 T G C20 A C A C A G T A - 3'
3' - G G A A G G C T C A G A G G G T G A C G T G T G T C A T - 5'
b
In vitro on-target validation in cell lines
Validation in liver organoids derived from Pahenu2 mice
In vivo correction of Pahenu2 mice
c
Per cent of total sequencing reads with C converted to T
50%
40%
30%
20%
10%
0
C2
C5
C6
C11
C13
C14
C15
C17
C20
****
*
ns
untreated
nSaCas9-BE3 + sgRNA
dLbCpf1-BE + crRNA
dLbCpf1-BE + crRNA + SpCas9 nickase
dLbCpf1-minBE + crRNA
d
Per cent of total sequencing reads
100%
80%
60%
40%
20%
0%
unedited
correctly edited
incorrectly edited
Indels
nSaKKH-BE3 + sgRNA
dLbCpf1-BE + crRNA
dLbCpf1-BE + crRNA + SpCas9 nickase
dLbCpf1-minBE + crRNA

Figure 1

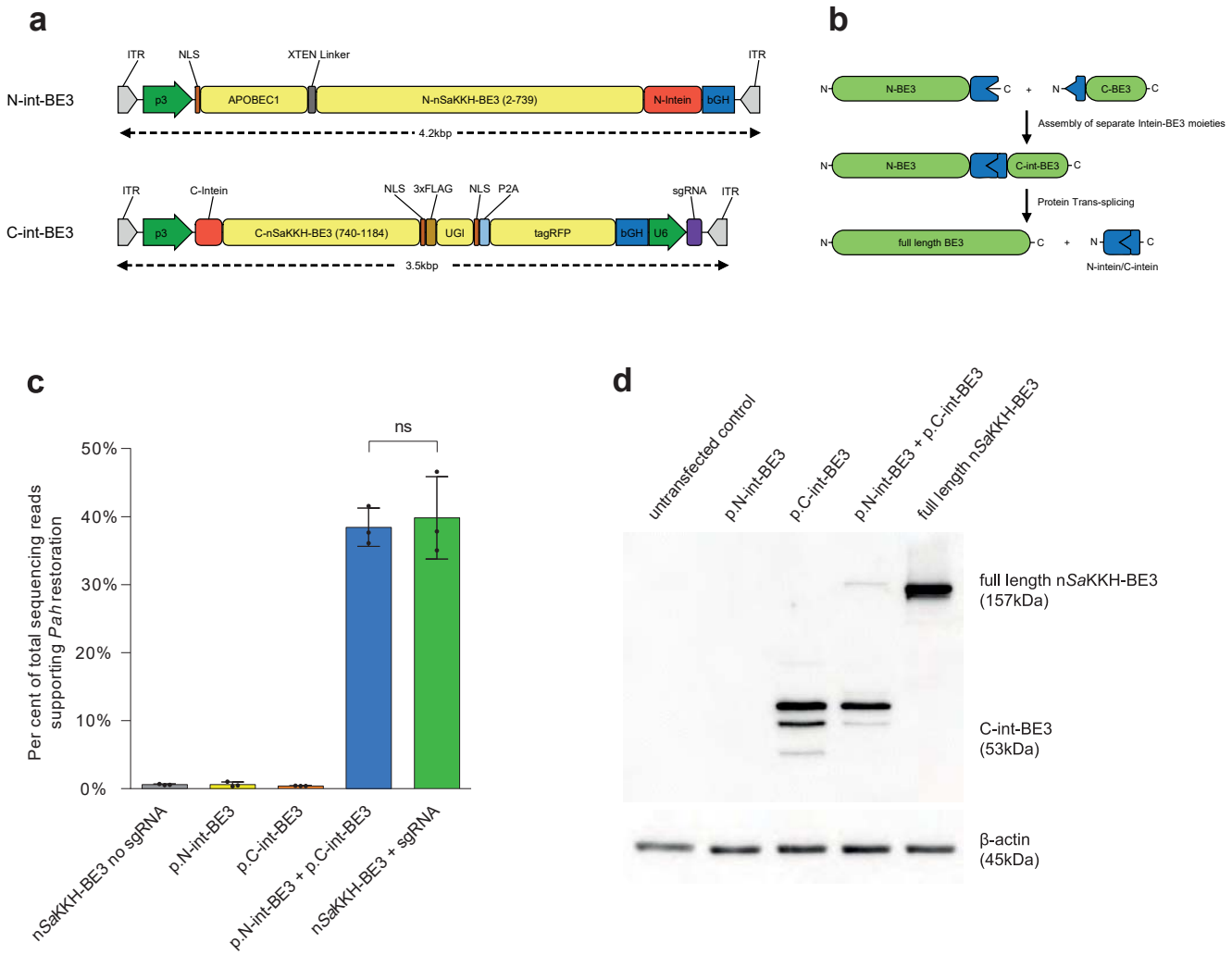

a
N-int-BE3
ITR
NLS
XTEN Linker
p3
APOBEC1
N-nSaKKH-BE3 (2-739)
N-Intein
bGH
ITR
4.2kbp
C-int-BE3
ITR
C-Intein
NLS
3xFLAG
NLS
P2A
sgRNA
ITR
p3
C-nSaKKH-BE3 (740-1184)
UGI
tagRFP
bGH
U6
3.5kbp
b
N-BE3
C
+
N
C-BE3
C
Assembly of separate Intein-BE3 moieties
N-BE3
C-int-BE3
Protein Trans-splicing
full length BE3
N-intein/C-intein
c
ns
Per cent of total sequencing reads supporting Pah restoration
50%
40%
30%
20%
10%
0%
nSaKKH-BE3 no sgRNA
p.N-int-BE3
p.C-int-BE3
p.N-int-BE3 + p.C-int-BE3
nSaKKH-BE3 + sgRNA
d
untransfected control
p.N-int-BE3
p.C-int-BE3
p.N-int-BE3 + p.C-int-BE3
full length nSaKKH-BE3
full length nSaKKH-BE3 (157kDa)
C-int-BE3 (53kDa)
β-actin (45kDa)

Figure 2

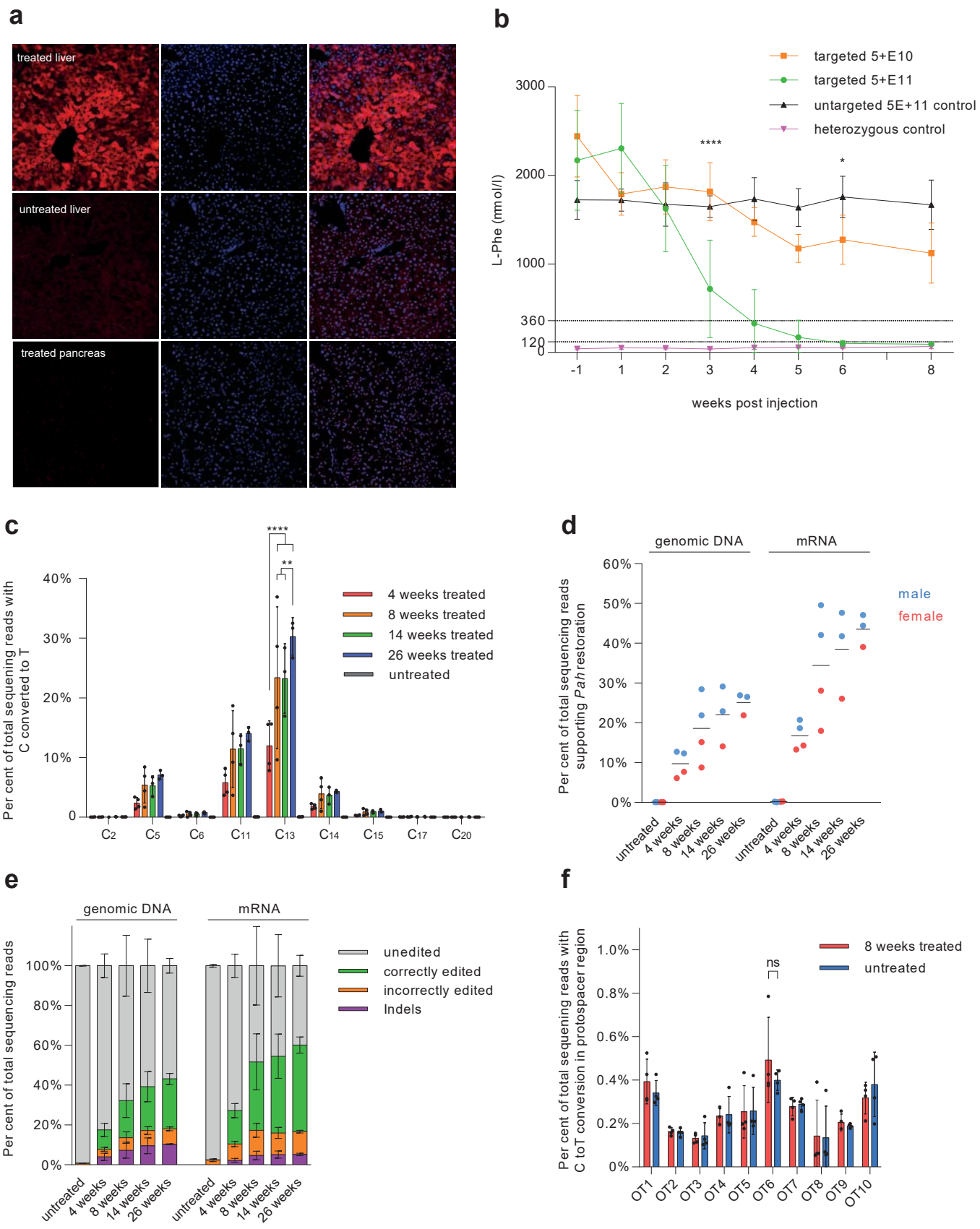
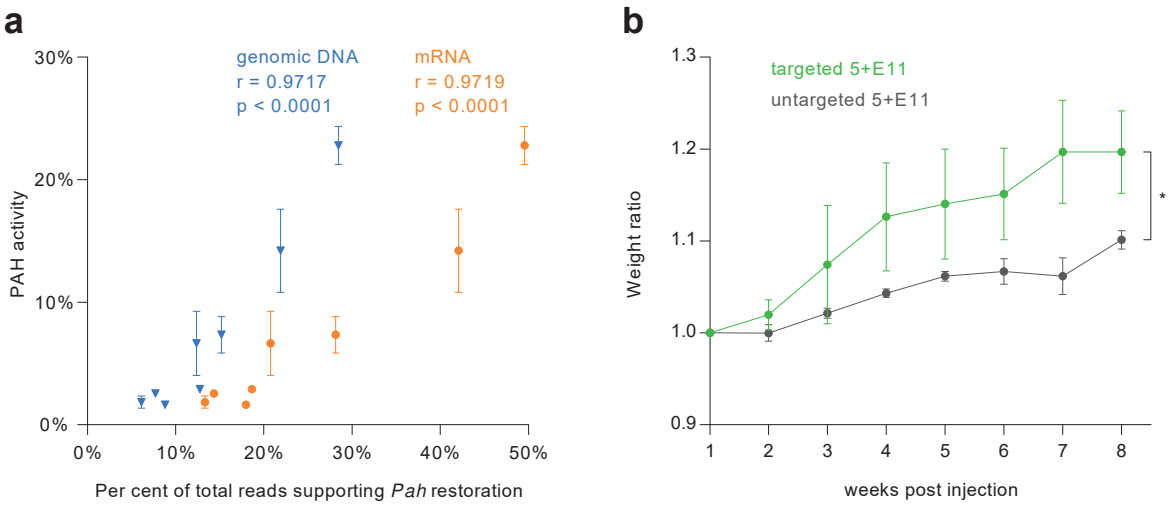

a
treated liver
untreated liver
treated pancreas
b
targeted 5+E10
targeted 5+E11
untargeted 5E+11 control
heterozygous control
L-Phe (mmol/l)
3000
2000
1000
360
120
0
****
*
-1
1
2
3
4
5
6
8
weeks post injection
c
Per cent of total sequening reads with C converted to T
40%
30%
20%
10%
0
****
**
4 weeks treated
8 weeks treated
14 weeks treated
26 weeks treated
untreated
C2
C5
C6
C11
C13
C14
C15
C17
C20
d
genomic DNA
mRNA
Per cent of total sequencing reads supporting Pah restoration
60%
50%
40%
30%
20%
10%
0%
male
female
untreated
4 weeks
8 weeks
14 weeks
26 weeks
e
genomic DNA
mRNA
Per cent of total sequencing reads
100%
80%
60%
40%
20%
0%
unedited
correctly edited
incorrectly edited
Indels
untreated
4 weeks
8 weeks
14 weeks
26 weeks
f
Per cent of total sequening reads with C toT conversion in protospacer region
1.0%
0.8%
0.6%
0.4%
0.2%
0%
ns
8 weeks treated
untreated
OT1
OT2
OT3
OT4
OT5
OT6
OT7
OT8
OT9
OT10

Figure 3



c

Pah^{enu2} rAAV8.N-int-BE3 + rAAV8.C.int-BE3.sgRNA

Pah^{enu2}

C57BL/6

Figure 4